Int J Life Cycle Assess (2008) 13:456–467
DOI 10.1007/s11367-008-0019-7

PANEL-BASED DECISION MAKING • DISCUSSION AND METHODOLOGY

Applying voting rules to panel-based decision making in LCA

Christoph Koffler · Liselotte Schebek · Stephan Krinke

Received: 14 April 2008 / Accepted: 21 May 2008 / Published online: 19 June 2008

Abstract
Background, aim, and scope Cross-category weighting is one possible way to facilitate internal decision making when dealing with ambiguous impact assessment results, with simple additive weighting being a commonly used method. Yet, the question as to whether the methods applied today can, in fact, identify the most "environmentally friendly" alternative from a group perspective remains unanswered. The aim of this paper is to propose a new method for group decision making that ensures the effective identification of the most preferable alternative.
Materials and methods Common approaches to deduce a single set of weighting factors for application in a group decision situation (e.g., arithmetic mean, consensus) are discussed based on simple mathematics, empirical data, and thought experiments. After proposing an extended definition for "effectiveness" in group decision making, the paper recommends the use of social choice theory whose main focus is to identify the most preferable alternative based on individuals' rankings of alternatives. The procedure is further supplemented by a Monte Carlo analysis to facilitate the assessment of the result's robustness.
Results The general feasibility of the method is demonstrated. It generates a complete ranking of alternatives, which does not contain cardinal single scores. In terms of effectiveness, the mathematical structure of the procedure ensures the eligibility for compromise of the group decision proposal. The sensitivity analysis supports the decision makers in understanding the robustness of the proposed group ranking.
Discussion The method is based upon an extended definition of effectiveness which acknowledges the eligibility for compromise as the core requirement in group decision contexts. It is shown that multi-attribute decision-making (MADM) methods in use in life cycle assessment (LCA) today do not necessarily meet this requirement because of their mathematical structure. Further research should focus on empirical proof that the generated group results are indeed more eligible for compromise than results generated by current methods that utilize an averaged group weighting set. This is closely related to the question considering under which mathematical constraints it is even possible to generate an essentially different result.
Conclusions The paper describes a new multi-attribute group decision support system (MGDSS) for the identification of the most preferable alternative(s) for use in panel-based LCA studies. The main novelty is that it refrains from deducing a single set of weighting factors which is supposed to represent the panel as a whole. Instead, it applies voting rules that stem from social choice theory. Because of its mathematical structure, the procedure is deemed superior to common approaches in terms of its effectiveness.
Recommendations and perspectives The described method may be recommended for use in internal, panel-based LCA studies. In addition, the basic approach of the method—the combination of MADM methods with social choice theory—can be recommended for use in all those situations where multi-attribute decisions are to be made in a group context.

Keywords Group decision support system · Multi-attribute decision making · Panel methods · Social choice theory · Voting · Weighting

Responsible editor: David Hunkeler

C. Koffler (✉) · S. Krinke
Volkswagen AG, Environmental Affairs Product,
P.O. Box 011/1774, 38426 Wolfsburg, Germany
e-mail: christoph.koffler@volkswagen.de

L. Schebek
Forschungszentrum Karlsruhe,
Institute for Technology Assessment and System Analysis,
Department of Technology-Induced Material Flows,
Hermann-von-Helmholtz-Platz 1,
76344 Eggenstein-Leopoldshafen, Germany

1 Background, aim, and scope

When applying life cycle assessment (LCA) to decision making in industrial eco-design, it is desirable to provide the respective decision makers with preferably unambiguous conclusions regarding the environmental superiority of certain alternatives over others (Goedkoop and Spriensma 2000). However, LCA practitioners frequently face situations where there is no single alternative which performs best in all of the considered impact categories. Cross-category weighting is one possible way to facilitate internal decision making while dealing with such "trade-offs" and is consistent with ISO requirements as long as the results of such comparative assertions are not disclosed to the public (International Organization for Standardization 2006). Several authors have pointed out that this approach bares close resemblance to multi-attribute decision making (MADM) (Benoit and Rousseaux 2003; Hertwich and Hammitt 2001; Lundie 1999; Mettier et al. 2006; Seppälä et al. 2002). Simple additive weighting (SAW) is a less demanding and therefore a commonly used MADM method in LCA today (Bovea et al. 2007a, b; Finnveden et al. 2006; Finnveden et al. 2002; Garrido and Alvarez del Castillo 2007; Goedkoop and Spriensma 2001; Güereca et al. 2007; Kicherer et al. 2007; Lassaux et al. 2007; Lundie 1999; Perzon et al. 2007; Rex and Baumann 2007; Yusoff and Hansen 2007). In addition, panel methods have been recognized as being one of the most promising ways to elicit the required category weighting factors (Finnveden et al. 2002). At the same time, important business decisions regarding design alternatives, investments, etc. are likely to be made by groups rather than by a single decision maker (Geldermann 2006). Yet, little attention has been given to the fact that most multi-attribute decision making (MADM) methods originally have been developed to only support a single decision maker (Lundie 1999). So, the question as to whether these methods are indeed effective in identifying the most "environmentally friendly" alternative from a group perspective remains unanswered. Accordingly, Hertwich and Hammitt (2001) state that the effectiveness of decision support methods is "largely unexplored in LCA" and (Hofstetter 1996) points out that, although decision makers agree with the use of multi-attribute decision making methods, they frequently disagree with the outcomes.

2 Life cycle assessment and group decision making

Hwang and Lin (1987) offer a formal approach to group decision making under multiple criteria for the selection of alternatives. Accordingly, in the case of life cycle impact assessment based on impact categories, the evaluation of alternatives can be characterized as follows:

- The employed criteria in the form of category indicator results are cardinal measures, e.g., 12 t of CO_2-equivalents.
- The scale transformation is performed via normalization to a reference system, e.g., as in Western Europe.
- The final normalization to an interval [0; 1] is achieved by dividing all values by the maximum value.
- All group members apply the same set of agreed criteria; and thus, the same set of impact categories.

The solution to such a group decision problem can be formulated as a set of matrices, where m alternatives (i=1,..,m) are evaluated by p group members (k=1,.., p) via n impact categories (j=1,.., n):

$$R^k = \left[r_{ij}\right]^k = \begin{bmatrix} r_{11} & .. & r_{1j} & .. & r_{1n} \\ r_{21} & .. & r_{2j} & .. & r_{2n} \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ r_{m1} & .. & r_{mj} & .. & r_{mn} \end{bmatrix} \quad (1)$$

The evaluation process can then be summarized by a mapping function ψ that leads to a group decision and can be obtained through ranking, rating, scoring, or voting (Hwang and Lin 1987):

$$\psi : \left\{R^k | k = 1, .., p\right\} \rightarrow \{G\} \quad (2)$$

Another important characteristic of group decision processes is whether they possess a hierarchic or non-hierarchic structure (Vetschera 1991). A hierarchic structure means that individual processes and group processes take place one after the other, with initial evaluation on an individual level preceding the subsequent aggregation into a group result. This can be accepted as being the case for panel methods, as weight elicitation is usually an individual process that is performed without direct interaction between panel members (Lundie and Huppes 1999; Mettier et al. 2006; Soares et al. 2005). Also, the hierarchic structure is predetermined by the temporal and/or spatial separation of the panel members in cases of elicitation via questionnaires.

There are two basic group decision models of hierarchic systems (Vetschera 1991). In the first one, denoted (a) in Fig. 1, individual evaluations are in some way combined with the respective data base to arrive at a group result. No individual results are generated along the way. The other possible way of deducing a group result, denoted (b) in Fig. 1, is to first calculate individual results based on individual evaluations and to then aggregate these into a group result.

Panel methods have been identified as one of the most promising approaches to support decision making in LCA and are widely preferred by the LCA community (Finnveden

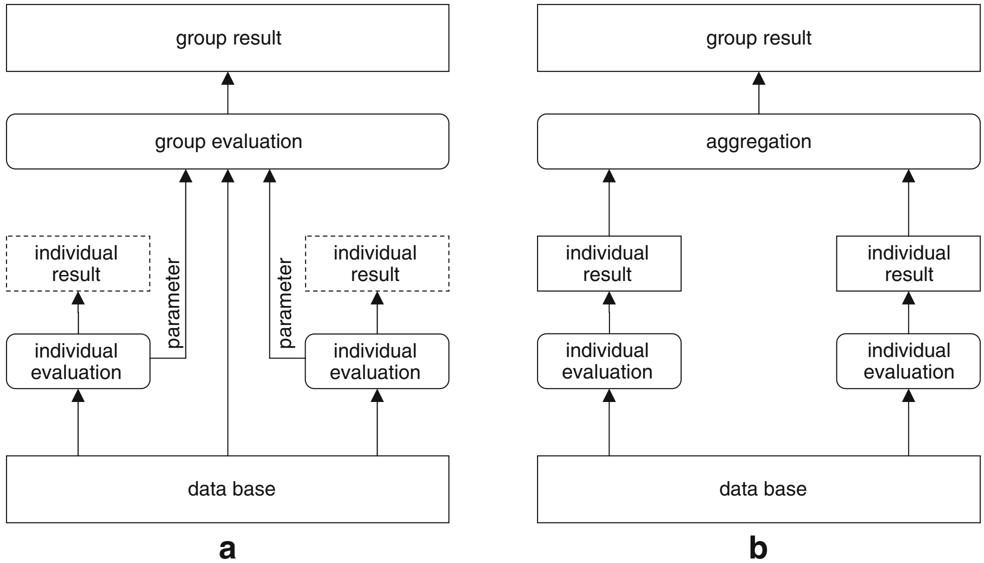

Fig. 1 Group decision models of hierarchic systems (Vetschera 1991)

et al. 2002; Soares et al. 2005). The commonly used simple additive weighting (SAW) is a multi-attribute decision-making method that assigns a panel weighting factor w_j to each impact category. Subsequently, it calculates one single score for each alternative by means of a weighted sum (Lundie 1999):

$$V_i = \sum_{j=1}^{n} w_j * r_{ij} \tag{3}$$

V_i single score of alternative A_i
w_j weighting factor of impact category j with $\sum w_j = 1 (j = 1,..,n)$
r_{ij} indicator result of alternative i in impact category j

In practice, the respective panel-weighting factor w_j is obtained by aggregating the elicited individual weights of each impact category either via compromise or mathematical models (Soares et al. 2005). Therefore, the application of panel methods to LCA can be characterized as a hierarchic group decision-making process that corresponds to the group decision model depicted in part (a) of Fig. 1. Next, the appropriateness of this approach in terms of effectiveness will be discussed.

3 Discussion of the effectiveness of common approaches

3.1 Effectiveness in group decision making

Before discussing the effectiveness of common approaches to decision making in panel methods, one will first have to define "effectiveness" in the given context. In the context of comparative LCA studies, the objective clearly is the identification of the most "environmentally friendly" product, process, or service. Because the term "environmentally friendly" is value laden itself (Steen 2006), the identification of the best alternative in the presence of trade-offs becomes a value choice. Lundie (1999) defines the most preferable alternative in LCA as the one that "causes the least environmental impact depending on the underlying value system. The most preferable alternative is the best possible representation of the decision-maker's requirements. These requirements reflect the decision-maker's objectives." Note that this definition addresses a single decision maker, according to the scope of most MADM methods. In group situations though, the final result should be as eligible for compromise as possible (Geldermann 2006). The closer a decision-making procedure approaches a hypothetical consensus, where all panel members agree on the same alternative as being the most preferable, the more eligible for compromise the generated result is. The above definition of the most preferable alternative in LCA therefore needs to be extended to group decision making as follows:

The most preferable alternative causes the least environmental impact depending on the underlying individual value systems. It poses a recommendation for compromise, which is the best possible representation of all decision makers' requirements. Such a compromise requires the approximation to an ideal consensus, where there is unanimity about the result.

3.2 Aggregation by arithmetic mean

As mentioned above, it is a common approach to deduce a group weighting set by calculating the arithmetic mean of all individual weighting factors per impact category (Mettier et al. 2006). Yet, it is also well-known that these individual weighting factors tend to differ considerably between individuals and to spread over large bandwidths,

which is due to the subjective nature of value choices (Lundie 1999; Lundie and Huppes 1999; Soares et al. 2005; Stahl 1999). While statistic measures generally help to better understand any survey's results, their consideration in the subsequent weighting process is limited, as additive MADM models such as SAW require a single weight per impact category. Now, the main problem is that the arithmetic mean does not contain any information about spread or distribution (Lundie and Huppes 1999). Also, the arithmetic mean is calculated without taking into account the weights that the respective panel member assigned to other impact categories. This means that the arithmetic mean is blind to the individual's preferences towards other impact categories, which is of high significance in the case of cross-category weighting. If effectiveness depends on eligibility for compromise, the objective is to then meet the individual preferences of as many panel members as possible. It is therefore advisable to preserve and carefully regard the individual weights applied to each impact category. For this reason, and because of the general inability to consider spread and distribution, one should abstain from applying the arithmetic mean in group decision-making processes.

3.3 Aggregation based on individual results

One possible way to preserve each panel member's preferences in the decision-making process is to calculate a single score for each alternative and each panel member before aggregating it into a group result. This corresponds to the procedure depicted in part (b) of Fig. 1. The intermediate result following Eqs. 1 and 3 is a matrix with p columns and m rows:

$$[V_{ik}] = \begin{bmatrix} V_{11} & .. & V_{1k} & .. & V_{1p} \\ V_{21} & .. & V_{2k} & .. & V_{2p} \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ V_{m1} & .. & V_{mk} & .. & V_{mp} \end{bmatrix} \quad (4)$$

The task, therefore, is to deduce a group result based on individual single scores per alternative. This corresponds to the structure of multi-attribute group decision making as described by Geldermann (2006). Accordingly, because MADM methods such as SAW belong to the class of aggregation models, it is permitted to aggregate their results via the same or another MADM method. In the case of SAW, this means the (weighted) summation of all individual single scores per alternative. Table 1 shows a fictitious example of three alternatives A, B, and C that are evaluated by a panel of six decision makers. Their individual single scores are summed up for each impact category to arrive at a group result $B \succ A \succ C$. No weights are applied during this last summation, as each panel member's opinion is accepted as being equally significant to the group result.

Table 1 Example of a conflict between group decision and eligibility for compromise

	A	B	C	Rank order
Person 1	0.3	0.5	0.7	$A \succ B \succ C$
Person 2	0.8	0.5	1.0	$B \succ A \succ C$
Person 3	0.5	0.3	0.2	$C \succ B \succ A$
Person 4	0.1	0.3	0.2	$A \succ C \succ B$
Person 5	0.5	0.1	0.3	$B \succ C \succ A$
Person 6	0.9	1.0	0.8	$C \succ A \succ B$
Panel (sum)	3.1	2.7	3.2	$B \succ A \succ C$

The displayed values denote resulting single scores from SAW
'$\succ$' denotes the logical connection 'preferred to'
'$\sim$' would denote indifference between alternatives

This approach also corresponds to welfare economics, where societal utility is assumed to be the sum of all individual utilities (Tangian 2000). Nevertheless, Table 1 contains a conflict between group decision and eligibility for compromise. If the rank frequencies of the alternatives are taken into account, it becomes apparent that each alternative holds a first rank twice in the individual rankings, a second rank twice, and a third rank twice. This, in turn, means that four out of six panel members would have to disagree with the group result, which appoints alternative B to be the most preferable one. Under the condition that all panel members exhibit a rational behavior, which is one of the basic assumptions of decision–theoretic aggregation models (Hwang and Lin 1987), the panel should therefore reject B as being the most preferable alternative. This conflict is again a consequence of neglecting available rank information in the aggregation process, where two identical values may carry different meanings. The careful consideration of rank information appears to be crucial to the eligibility for compromise and cannot be disregarded.

3.4 Aggregation by compromise

To avoid the problems described so far, it would be helpful to first establish a compromise about a group weighting set before cross-category weighting is performed. There are basically two ways to do so in a panel process: through mediation or the Delphi method (Braunschweig et al. 1996). Mediation means that a qualified mediator tries to establish a compromise in an open discussion. The Delphi method is a cyclical, questionnaire-based survey method, where experts are repeatedly asked to reconsider their evaluations based on the results of the previous evaluation. Both methods are apt to reduce both bandwidth and spread of the weighting factors. Yet, the larger the number of panelists, the more divers their background and the more numerous the impact categories, the less likely it becomes that a panel can agree on a single weighting set. This is

because of the heterogeneity of the panel member's risk attitudes and information structures (Geldermann 2006). Although both the mediation and Delphi method could in theory be continued until a compromise is reached, the individual motives for consent remain undisclosed. There would be no formal proof that a so-enforced compromise weighting set optimally represents all panel members' preferences. Especially in mediation, there are problems concerning group dynamics such as dominance of some panel members, peer pressure, or deficiencies of articulation (Geldermann 2006). In addition, Delphi methods suffer from an unfavorable decrease in participation from one round to the next. If eligibility for compromise is the main objective, one should therefore also refrain from applying compromise weighting sets in the decision-making process.

Based on the definition given in Section 3.1, one can therefore conclude that the described common ways of aggregating individual preferences into a group perspective do not necessarily fulfill the requirements for an effective group decision making. This is the case either because they neglect available rank information or because they are based on rather intransparent processes, where the individual degree of consent with a compromise weighting set remains undisclosed.

4 Towards a new approach to group decision support in LCA

4.1 Social choice theory

It has been shown that common approaches to cross-category weighting do not inevitably integrate individual preferences in the decision-making process in an effective way. To ensure the acceptance of the group result by the majority of the panel members, it appears to be necessary to preserve each panel member's preferences throughout the procedure, and to accurately regard the resulting rank information as well. Formally, one is looking for a mapping function ψ (Eq. 2) that is capable of achieving this goal based on individual rankings of alternatives (Eq. 4).

The identification of the most preferable "candidate" based on individual rank orders has been the main focus of social choice theory for the last two centuries (Tideman 1987). It is concerned with the most democratic representation of the people's will in political elections. Social choice theory therefore appears especially capable of ensuring the eligibility for compromise. Social choice theory knows more than 25 different voting rules (Tideman 1987; Vandercruyssen 1999), which can be subdivided into two main classes: weighted scoring rules and Condorcet voting rules. While weighted scoring rules assign a certain score to each rank in every individual ranking, Condorcet voting rules are based on pairwise comparisons and make use of the existing majorities to deduce a group result (Condorcet 1785; Gehrlein 1998). Subsequently, common voting rules are discussed concerning their appropriateness in the given context.

4.2 Requirements and method selection

Condorcet consistency The Marquis de Condorcet (1743–1794), one of the founders of social choice theory, demanded that if there is a candidate A that wins all pairwise comparison against his contestants, then this so-called "Condorcet winner" should be elected winner of the ballot (Condorcet 1785). A Condorcet-consistent voting rule will always select the Condorcet winner if one exists. This requirement appears to be both logical and fair because, if a voting rule for some reason elected another candidate, then, the Condorcet winner should prevail in a runoff election between these two (Green-Armytage 2004). This criterion therefore appears to be relevant in the given context. Although the aforementioned, weighted scoring rules (WSR) may select the Condorcet winner in some cases, they do not necessarily select them in all cases (Vandercruyssen 1999) and will therefore be excluded from further method selections.

Consideration of cardinal information As displayed in Table 1, there is cardinal information about distances between ranks in the form of single scores in LCA, which expresses the individual strength of preferences towards the alternatives. However, most Condorcet voting rules assume that the voter solely provides ordinal rankings. Green-Armytage (2004) proposed the only Condorcet-consistent voting rule that demands that the voters not only provide ordinal rankings of all candidates, but also assigns cardinal utility measures to each candidate. While this approach poses quite a challenge in real elections, it superbly matches the data at hand in LCA. The so-called "cardinal weighted pairwise comparison" (short: cardinal pairwise) calculates the strength of defeat S per pairwise comparison of two alternatives, A_i and A_j, by summing up the differences between the alternatives' cardinal measures over all those individual rankings that form the majority opinion:

$$S\left(A_i \succ A_j\right) = \sum_{k=1}^{p^*} \left(V_{jk} - V_{ik}\right) \text{ with } p^* = \left\{k \middle| A_i \succ A_j\right\}, \quad i \neq j \text{ and } V_{ik} < V_{jk} \tag{5}$$

For instance, if there are six people in a panel of ten that prefer A to B ($A \succ B$), then, the differences between the cardinal measures of A and B are summed up over their respective six rankings. Subsequently, the pairwise state-

ments may be aggregated by use of any Condorcet voting rule. Green-Armytage (2004) recommends the methods of Schulze and the ranked pairs rule among others (Schulze 2003, 2008; Tideman 1987; Zavist and Tideman 1989). These two are deemed the “most supportable” Condorcet-consistent voting rules available today (Johnson and Stahl 2006).

Approximation to a hypothetical consensus For eight voting rules (WSR and Condorcet), Meskanen and Nurmi (2005, 2006) showed that they each try to approximate a hypothetical consensus as close as possible in their own way. This is in line with the definition of effectiveness of group decisions as proposed before. Only this kind of voting rules should therefore be considered for use in LCA. The method of Schulze and ranked pairs both fulfill this requirement.

Complexity of algorithm The more sophisticated a procedure, the less transparent and trustworthy the outcome becomes from the decision makers’ point of view (Schuh 2001). A fundamental requirement is that the decision makers have to be able to understand the basic operations performed by the respective group decision support system. Because its complexity is the main objection against the method of Schulze (Johnson and Stahl 2006), it will therefore be disregarded for use in LCA.

Based on the described requirements and available state-of-the-art voting rules, a combination of cardinal weighted pairwise comparison and ranked pairs rule is selected. Subsequently, the basic procedure and necessary adaptions to the specific requirements of decision making in LCA will be explained.

5 Method description and modification

5.1 Cardinal pairwise and ranked pairs rule

The cardinal weighted pairwise comparison calculates direction and strength for each pairwise defeat via Eq. 5. The ranked pairs rule then sorts all pairwise statements in descending order based on their strengths of defeat. The order in which the pairwise comparisons are taken into account is of relevance to the total outcome if and only if there is a logical inconsistency in the panel’s preference structure. For instance, if there is a majority for $A \succ B$, $B \succ C$, and $C \succ A$, then this constitutes a so-called “Condorcet paradox” where $A \succ B \succ C \succ A$ (Kurrild-Klitgaard 2001; Tideman 1985). The ranked pairs rule avoids this conflict through the stepwise consideration of pairwise defeats in descending order of strength of defeat. Subsequent pairwise defeats that cause a logical conflict are neglected (Tideman 1987; Zavist and Tideman 1989).

5.2 Method modification for use in LCA

Although the utilization of existing voting rules seems appropriate to enhance the eligibility for compromise of a group result, the general scope of social choice theory is not fully compliant with that of LCA. The procedure therefore needs to be modified to the requirements of decision making in LCA. The implemented modifications are described below.

Permissibility of indifference between alternatives When talking about social choice theory and political elections, in general, it is obvious that the goal is to appoint a single winner, e.g., the next president, mayor, chairman, or the like. In LCA, it is not reasonable to limit the process to a single winner. If there are two or more alternatives that appear to be equally preferable from the panel’s perspective, then this should be reflected by the group result as well. It therefore becomes permissible to state not only that $A \succ B$ or $B \succ A$, but also that $A \sim B$ (indifference). This specification is closely connected to the question of the underlying majority. Conventional Condorcet voting rules demand a simple majority of >50% and disregard all panel members’ rankings that display indifference between the respective alternatives. If such indifference becomes permissible for consideration in the group result, then the decision support system has to extend the pairwise comparisons to indifference statements as well. This, in turn, requires the employment of an absolute instead of a simple majority. Under absolute majority, all pairwise statements that are supported by more than 50% of the panel members will be accepted for consideration in the group result. If there is not any pairwise statement that is supported by more than 50% of the panel, then the statement ‘~’ (indifference) is processed.

Strength of disagreement as sorting criterion If one accepts indifference statements for consideration in the group result, then there is the need for a modified sorting criterion, as Eq. 5 calculates the strength of defeat only in the direction of the defeat. For indifference statements, there is no such direction, hence $S(A \sim B) = 0$. If one interprets the ranked pairs rule as a mean of maximizing the support within the panel, then one may also consider a sorting criterion that reversely tries to minimize the existing disagreement. We therefore propose a sorting criterion $\overline{S}$ (strength of disagreement) that calculates the sum of all differences between the single scores of the respective alternatives over all those rankings that do not support the majority opinion:

$$\overline{S}(A_i \succ A_j) = \sum_{k=1}^{p^*} \left(V_{ik} - V_{jk}\right) \text{ with } p^* = \left\{k \middle| A_j \succ A_i\right\}, \quad i \neq j \text{ and } V_{ik} > V_{jk} \tag{6}$$

For an indifference statement $A_i \sim A_j$, the strength of disagreement $\overline{S}$ is then the sum of the support S for $A_i \succ A_j$ and $A_j \succ A_i$. Therefore, $\overline{S}$ $(A_i \sim A_j)$ is calculated as follows:

$$\begin{aligned}\overline{S}(A_i \sim A_j) &= \overline{S}(A_i \succ A_j) + \overline{S}(A_j \succ A_i) \\ &= S(A_j \succ A_i) + S(A_i \succ A_j)\end{aligned} \tag{7}$$

as

$$\begin{aligned}\overline{S}(A_i \succ A_j) &= S(A_j \succ A_i) + \underbrace{S(A_i \sim A_j)}_{0} \\ &= S(A_j \succ A_i)\end{aligned} \tag{8}$$

The resulting procedure is divided into two parts. First, all pairwise comparison statements that are supported by an absolute majority are sorted in ascending order regarding the criterion $\overline{S}$. For example, if the statement $A \succ B$ receives an absolute majority while all other panel members are indifferent between these two, then $\overline{S}(A \succ B) = 0$. If there would be unanimous support in the panel for the statement $A \sim B$, then $S(A \sim B)$ would equal to zero as well. We therefore created a criterion that allows us to sort preference as well as indifference statements alike.

Secondly, the procedure turns to all those statements that fall short of an absolute majority and are counted as indifference between alternatives. These are again sorted in ascending order by means of Eq. 7. The deduction of a group result then takes into account the majority statements first and the non-majority statements afterwards. This ensures that minority opinions with large differences between individual single scores cannot dominate the group result, which also enhances the eligibility for compromise.

Qualified majority If eligibility for compromise is accepted as the core requirement of an effective group decision support system, it seems reasonable to base group evaluations about preference or indifference on a majority that is larger than the absolute majority of >50% of votes. For instance, if there are 8 people in a panel of 15 that support $A \succ B$, while seven support $B \succ A$, then $A \succ B$ prevails with the smallest possible advance of one supporter. Vice versa, a change in mind of only one panel member may reverse the preference statement. The procedure is therefore expanded to a qualified majority that can be specified by the user. The default is set to a two thirds majority. Again, all pairwise comparisons that fall short of this majority are counted as indifference. This approach tightens the requirements to be fulfilled for a preference statement and protects the group result from rank reversals because of minor changes through the installation of an indifference interval between $A \succ B$ and $B \succ A$.

Sensitivity analysis Sensitivity analyses are an essential part of both LCA and decision support systems (Geldermann 2006). A compromise is more likely to be accepted if the sensitivity of the proposed group result towards changes in the underlying data is communicated to the group members (Zhang 2004). The procedure is therefore supplemented by a Monte Carlo analysis that is capable of coupling LCIA results as well as individual weighting factors with uncertainty intervals.

6 Application to a case study

6.1 Selection of alternatives

The developed algorithm was applied to an internal study of three automotive parts that has been conducted within the Volkswagen Group Research Department in 2001. They are made anonymous (alternative A, B, and C), because the main objective of this case study is to demonstrate the functionality of the proposed procedure. This case study is selected because it fulfills two conditions for which the method is deemed especially beneficial: the comparison of more than two alternatives and the existence of a distinct trade-off between impact categories. None of the alternatives in the case study scores best in all impact categories (GWP100, AP, EP, and POCP) from cradle to grave. The selection of the most ‘environmentally friendly’ alternative therefore becomes a value choice (Finnveden et al. 2002; Schmidt and Sullivan 2002; Steen 2006). In addition, the simultaneous consideration of 12 numerical values (3 alternatives times 4 impact categories) poses a significant challenge to most people, as this is beyond their cognitive limits (Miller 1967). The employment of a mathematical decision support system is therefore highly recommended.

6.2 Elicitation of weighting factors

For the case study, the weighting factors per impact category were elicited among the environmental experts working for the Volkswagen AG (Koffler 2007). All employees whose day-to-day business has a clear emphasis on environmental issues were considered to be environmental experts in the survey. This assumption holds for roughly 150 employees worldwide. They were each contacted via e-mail and asked to state their personal preferences regarding ten different impact categories in a global context based on the question “Which environmental problem should preferably be diminished by the Volkswagen Group in your personal opinion?” This was preceded by a short description of the impact categories to establish a common level of knowledge. Following Finnveden et al.

(2002), a ratio estimation technique was used in the questionnaire. The respondents were asked to assign ten points to the most important category and one to ten points to all others. The participation in the survey was 56%.

Figure 2 shows the bandwidths and SDs of the survey for the four impact categories considered in the case study. The respective arithmetic means are displayed in brackets behind the impact categories' names. It can be seen that the values for each impact category spread over large bandwidths and show considerable SDs. This heterogeneity is because of the general subjectivity of value judgments and is a common characteristic of this type of surveys (Lundie and Huppes 1999; Soares et al. 2005).

6.3 Identification of the most preferable alternative

The developed algorithm was implemented in the form of a prototype software tool. To initiate the evaluation, the user has to provide a file containing the impact assessment results at hand and another file containing the respective weighting factors as elicited in the panel process. The general procedure then takes place in analogy to Fig. 1b, deducing individual rank orders through SAW and then aggregating these by means of a voting rule as described in Section 5.

For the data at hand, the tool generates the group ranking "$A \succ C \succ B$". Therefore, concept A is appointed the most preferable one from a group perspective based on the LCIA indicator results and the elicited weighting factors. This is somewhat surprising as alternative A displays the highest LCIA indicator results in all categories but POCP, where it outperforms its contestants by far. Apparently, this advantage cannot be offset by the disadvantages in the other categories from the panel's point of view.

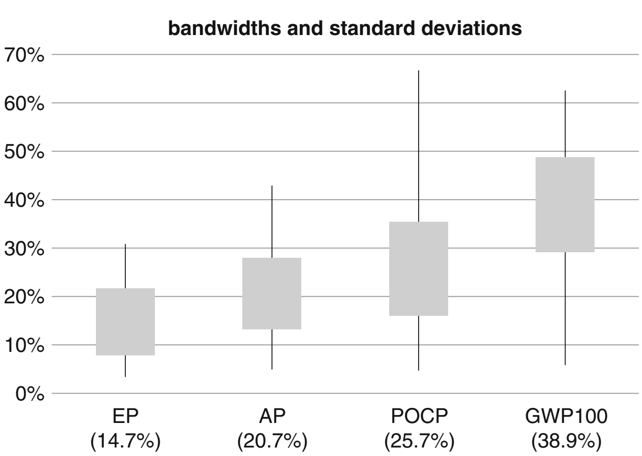

Fig. 2 Results of internal weight elicitation for application in case study

6.4 Sensitivity analysis

The algorithm is supplemented by a Monte Carlo analysis of the impact assessment results and the individual weighting factors to assess the robustness of the generated group result. Firstly, the user may specify a non-negative integer x, which is used to set up a uniformly distributed uncertainty interval ($\pm x$) to be applied to the elicited weighting factors. The application of this uncertainty interval to the utilized ten-point scale effectuates a reduction of the scale. For instance, if $x=1$, then the ten-point scale is effectively reduced to a four-point scale because 1 may mean 1 or 2, 2 may mean 1, 2, or 3, and so on. Only the value 10 is not associated with any uncertainty, as the task within the elicitation process was to assign this value to the most important impact category first. Therefore, this judgment shall not be questioned in the sensitivity analysis. An uncertainty interval larger than ±1 (default setting) is not considered useful in this context because it strongly diminishes the informational value of the weighting factors.

Secondly, the user may define a step width through a non-negative rational number from the interval [0; 1], which is used to set up uniformly distributed uncertainty intervals around the impact assessment results. These run between ±0% and 100%. This approach seems especially helpful if there are not any estimates available about the "true" uncertainties involved. If there are estimates available, then the tool also allows matching each normalized LCIA indicator result with a standard deviation. These may be obtained through prior Monte Carlo simulation within the respective LCA tool, though this was not done in the case study at hand.

In each run, the algorithm draws a random number from the interval around the impact assessment result for each alternative and impact category. These values are then combined with likewise randomly generated individual weighting factors to obtain a group result. This is reiterated as many times as specified by the user (default: $N=10{,}000$ runs). Then, the uncertainty interval is increased by the defined step width, and the algorithm repeats the procedure. After the maximal uncertainty interval of ±100% has been reached, a file containing the relative frequencies of all group rankings that appeared for each evaluated uncertainty interval is issued.

Figure 3 shows the results of a Monte Carlo analysis where the elicited weighting factors as well as the LCIA indicator results are subject to random selection from defined intervals as described above. The point of intersection where alternative A and C switch ranks is reached at an LCIA uncertainty interval of ±40%. One has to bear in mind that this uncertainty interval is capable to even out differences of close to 60% between any two LCIA

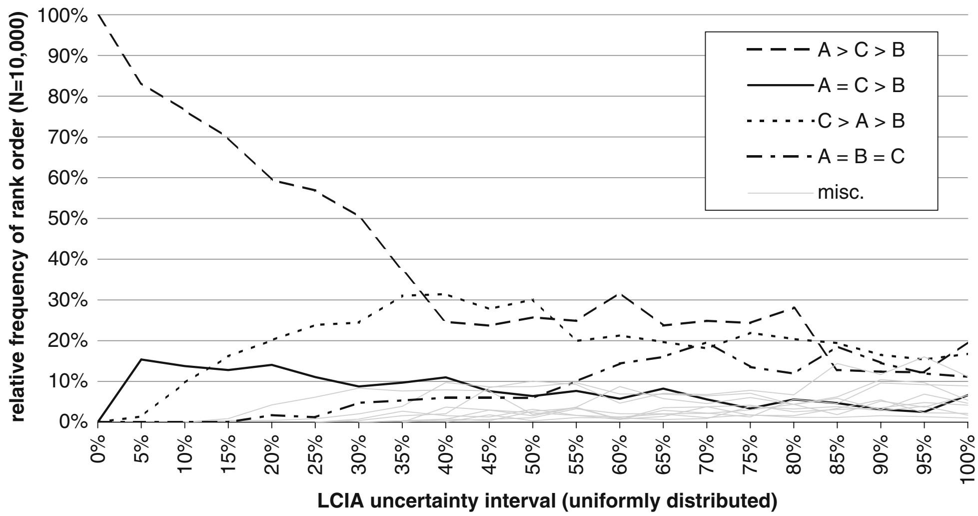

Fig. 3 Resulting rank order frequencies within sensitivity analysis

indicator results. This is a first hint at the original group ranking's robustness. Subsequently, the two alternatives switch ranks three more times within all considered uncertainties, which shows the closeness of the decision between the two alternatives when larger uncertainties are taken into account.

These results may be further discussed in two different ways. First, one may accept all LCIA uncertainty intervals between ±0% and 100% as being equally reasonable. This is the case when there is no information about the "true" uncertainties involved. Hence, one may also accept the cumulative frequency of each ranking as being decisive within the sensitivity analysis. For instance, the original group ranking "*A*≻*C*≻*B*" prevails in 39.2% of all 210,000 runs that were calculated. The second most frequent ranking is "*C*≻A≻B" with 19.3%. All other rankings prevail in less than 10% of all calculations. We therefore observe a substantial margin between first and second most frequent group rankings (Fig. 4).

Another way of discussing the sensitivity analysis' results is to only take into account those uncertainty intervals that are considered to be more meaningful than others. This may be based on expert judgments. For example, following (de Beaufort-Langeveld et al. 1997) differences between alternatives' impact assessment results that are below a certain percentage should be considered as equal. For instance, a difference of 20% between two impact indicator results can be bridged if an uncertainty interval larger than ±11.1% is applied to these values. In turn, one may argue that all rank orders that are based on smaller uncertainty intervals are not meaningful to the overall result and can be disregarded. Also, one may wish to limit the allowable maximum uncertainty to a value less than ±100%. For example, if there are reasons to believe

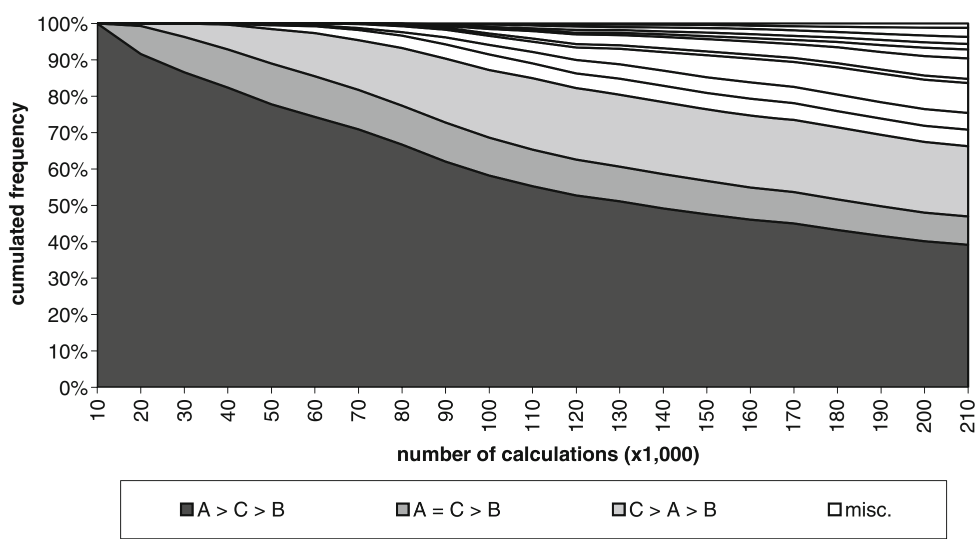

Fig. 4 Resulting cumulated rank order frequencies within sensitivity analysis

that differences of more than 50% between two impact assessment results should be preserved as being “different” throughout the calculation, then one should refrain from considering uncertainty intervals that are larger than ±33.3% in the sensitivity analysis.

Accordingly, if one only considers the sensitivity analysis results for uncertainty intervals between ±15% and 30% for the values at hand, then the original group ranking “$A \succ C \succ B$” prevails in 59.1% of all 40,000 runs, while the ranking “$C \succ A \succ B$” is rendered in 21.1% of these cases. Correspondingly, there are not any switches in rank order to be observed in Fig. 3 between these values. Hence, the consideration of only a defined section of all examined uncertainty intervals entails an even stronger margin in favor of the original ranking. The results also show that the maximum uncertainty interval of ±100% appears to be sufficient in this context. The larger the uncertainty intervals, the more numerous and less discriminate the resulting group rankings become.

Based on the described analyses, two main conclusions can be drawn regarding the case study:

- Alternative B is in any case the least environmentally friendly one from a group perspective.
- Although alternative C prevails under the assumption of some uncertainty intervals, the results of the sensitivity analysis show that the ranking “$A \succ C \succ B$” is robust. Alternative A is therefore elected the most “environmentally friendly” one from a group perspective based on the values at hand.

7 Method discussion

7.1 Benefits of the proposed method

The proposed method provides some notable benefits that shall be summarized in short. First of all, it preserves and carefully regards individual preferences throughout the decision-making process. Secondly, extreme positions of minorities cannot dominate the group result because the differences between individual single scores are only taken into account if the required majority is met. Thirdly, marginal differences between alternatives are less likely to affect the group result because pairwise comparisons that fall short of the qualified majority are processed as “equal”. Fourthly, the resulting rank order is less in need of interpretation as it does not contain cardinal measures such as conventional single scores that may display only marginal differences. Instead, it renders unambiguous preference statements such as “better” and “equal”. Lastly, the subsequent Monte Carlo simulation provides the respective decision makers with profound insight into the resulting group ranking’s sensitivity towards uncertainties. This kind of analysis can also be recommended for use in LCA studies that render a seemingly unambiguous result (e.g., dominance of one alternative) to test its robustness.

7.2 Extension to other aspects and fields of application

Just like any other multi-attribute decision-making (MADM) tool, the proposed method is not limited to the evaluation of life cycle impact assessment results. Relating to the “triple bottom line” of sustainability (Elkington 2002), an integration of other aspects such as costs or social indicators seems indicated to facilitate a more holistic evaluation of different design options. The integration of these aspects is possible following the theory of simple additive weighting (Fishburn 1967; Hwang and Yoon 1981). Also, a combination of the basic approach, the application of voting rules to decision making in group situations, with other MADM techniques such as MAUT/MAVT, AHP, TOPSIS, PROMETHEE, or ELECTRE is generally possible. In essence, the proposed method poses a multi-attribute group decision support system (MGDSS) that may be recommended for use in any multi-attribute group decision context that is based on cardinal measures, e.g., in product development, marketing, policy making and the like. In addition, the basic approach of combining MADM methods with social choice theory also applies to multi-attribute decisions based on ordinal measures, taking into account rank information only.

8 Summary and conclusions

The paper describes a new multi-attribute group decision support system (MGDSS) for the identification of the most preferable alternative/s for use in panel-based LCA studies. The main novelty is that it refrains from deducing a single set of weighting factors which is supposed to represent the panel as a whole. Instead, it applies voting rules that stem from social choice theory. In terms of effectiveness, the mathematical structure of the procedure ensures the eligibility for compromise of the group decision proposal. It is therefore deemed superior to common approaches of group decision making in LCA because these do not necessarily fulfill this requirement. They disregard available rank information or employ rather intransparent methods of deducting a compromise group weighting set.

According to its objective and procedure, the method is termed “clearLCA”, which stands for “consensus-led and environmentally appropriate recommendations”. It is supplemented by a Monte Carlo simulation that evaluates the robustness of the group result. The general feasibility of the method was demonstrated through its application to a case

study. The basic approach of the method—the combination of MADM methods with social choice theory—may further be recommended for use in all those situations where multi-attribute decisions are to be made in a group context.

9 Recommendations and perspectives

On the one hand, there should be empirical proof that the results generated by the proposed method are indeed more eligible for compromise than results generated by current methods. This, of course, is closely related to the question under which mathematical constraints is it even possible to generate an essentially different result. Future research should therefore focus on these issues.

In addition, the described way of considering uncertainties through Monte Carlo simulation may be complemented by more sophisticated methods to analyze the resulting rank frequencies. For instance, the question remains how large the margin between first and second most frequent group ranking ought to be to state that the most frequent one is the one to be deemed appropriate.

References

Benoit V, Rousseaux P (2003) Aid for aggregating the impacts in life cycle assessment. Int J Life Cycle Assess 8:74–82

Bovea M-D, Cabello R, Querol D (2007a) Comparative life cycle assessment of commonly used refrigerants in commercial refrigeration systems. Int J Life Cycle Assess 12:299–307

Bovea M-D, Saura Ú, Ferrero JL, Giner J (2007b) Cradle-to-gate study of red clay for use in the ceramic industry. Int J Life Cycle Assess 12:439–447

Braunschweig A, Förster R, Hofstetter P, Müller-Wenk R (1996) Developments on LCA valuation. Institut für Wirtschaft und Ökologie, Universität St. Gallen (IWÖ-HSG), St. Gallen

Condorcet Nd (1785) Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Reprint: Chelsea Publisher Co, New York, 1980

de Beaufort-Langeveld A, van den Berg NW, Haydock R, ten Houten M, Kotaji S, Oerlemans E, Schmidt W-P, Stranddorf HK, Weidenhaupt A (1997) Simplifying LCA—just a cut? Final report of the SETAC Europe LCA screening and streamlining working group. In: Christiansen K (ed) SETAC, Brussels

Elkington J (2002) Cannibals with forks: the triple bottom line of 21st century business. Capstone, Oxford XIV, 410 S pp

Finnveden G, Hofstetter P, Bare JC, Basson L, Ciroth A, Mettier T, Seppälä J, Johansson J, Norris GA, Volkwein S (2002) Normalisation, grouping, and weighting in life-cycle assessment. In: UdodeHaes HA et al (ed) Life cycle impact assessment—striving towards best practice. SETAC, Pensacola

Finnveden G, Eldh P, Johansson J (2006) Weighting in LCA based on ecotaxes—development of a mid-point method and experiences from case studies. Int J Life Cycle Assess 11:81–88

Fishburn PC (1967) Additive utilities with incomplete product set: applications to priorities and assignments. ORSA Publication, Baltimore

Garrido N, Alvarez del Castillo MD (2007) Environmental evaluation of single-use and reusable cups. Int J Life Cycle Assess 12:252–256

Gehrlein WV (1998) The sensitivity of weight selection on the condorcet efficiency of weighted scoring rules. Soc Choice Welfare 15:351–358

Geldermann J (2006) Mehrzielentscheidungen in der industriellen Produktion. Universitätsverlag Karlsruhe

Goedkoop M, Spriensma R (2000) Communicating LCA results—pragmatic tools and methods. Int Conf and exhibition on life cycle assessment—tools for sustainability. American Center for Life Cycle Assessment (ACLCA), Arlington

Goedkoop M, Spriensma R (2001) The eco-indicator 99—a damage oriented method for life cycle impact assessment. Methodology report. Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer, Den Haag

Green-Armytage J (2004) Cardinal-weighted pairwise comparison. Voting Matt 19:10–17

Güereca LP, Agell N, Gassó S, Baldasano JM (2007) Fuzzy approach to life cycle assessment—an application for biowaste management systems. Int J Life Cycle Assess 12:488–496

Hertwich EG, Hammitt JK (2001) A decision-analytic framework for impact assessment, part I—LCA and decision analysis. Int J Life Cycle Assess 6:5–12

Hofstetter P (1996) Towards a structured aggregation procedure. In: Braunschweig A, Förster R, Hofstetter P, Müller-Wenk R (eds) Development in LCA valuation. Institut für Wirtschaft und Ökologie, Eidgenössische Technische Hochschule (ETH), Zurich

Hwang C-L, Lin M-J (1987) Group decision making under multiple criteria: methods and applications. Lecture notes in economics and mathematical systems 281. Springer-Verlag

Hwang C-L, Yoon K (1981) Multiple attribute decision making: methods and applications—a state-of-the-art-survey. Lecture notes in economics and mathematical systems 186. Springer-Verlag, X, 259 pp

International Organization for Standardization (2006) ISO 14040: Environmental management—life cycle assessment—principles and framework. ISO, Geneva

Johnson PE, Stahl S (2006) Understanding modern mathematics. Jones & Bartlett Publishers, Sudbury

Kicherer A, Schaltegger S, Tschochohei H, Ferreira Pozo B (2007) Eco-efficiency—combining life cycle assessment and life cycle costs via normalization. Int J Life Cycle Assess 12:537–543

Koffler C (2007) Automobile Produkt-Ökobilanzierung. Institut WAR—Institut für Wasserversorgung und Grundwasserschutz, Abwassertechnik, Abfalltechnik, Industrielle Stoffkreisläufe, Umwelt- und Raumplanung. Technische Universität Darmstadt

Kurrild-Klitgaard P (2001) An empirical example of the condorcet paradox of voting in a large electorate. Public Choice 107:135–145

Lassaux S, Renzoni R, Germain A (2007) Life cycle assessment of water from the pumping station to the wastewater treatment plant. Int J Life Cycle Assess 12:118–126

Lundie S (1999) Ökobilanzierung und Entscheidungstheorie—praxisorientierte Produktbewertung auf der Basis gesellschaftlicher Werthaltungen. Springer-Verlag

Lundie S, Huppes G (1999) Product Assessment based on a range of societal preferences, global competitiveness through cleaner production. Australian Cleaner Production Association (ACPA), Brisbane, pp 441–451

Meskanen T, Nurmi H (2005) Measuring distance from consensus under various formulations of the social choice problem. Group Decision and Negotiation (GDN), Vienna

Meskanen T, Nurmi H (2006) Distance from consensus: a theme and variations, mathematics and democracy. Springer-Verlag

Mettier T, Scholz RW, Tietje O (2006) Measuring preferences on environmental damages in LCIA. Int J Life Cycle Assess 11:394–402

Miller GA (1967) The magic number seven, plus or minus two: some limits on our capacity for processing information. In: Alexis M, Wilson CZ (eds) Organizational decision making. Prentice-Hall

Perzon M, Johansson K, Fröling M (2007) Life cycle assessment of district heat distribution in suburban areas using PEX pipes insulated with expanded polystyrene. Int J Life Cycle Assess 12:317–327

Rex ELC, Baumann H (2007) Individual adaptation of industry LCA practice—results from two case studies in the Swedish forestry products industry. Int J Life Cycle Assess 12:266–271

Schmidt W-P, Sullivan JL (2002) Weighting in life cycle assessments in a global context. Int J Life Cycle Assess 7:5–10

Schuh H (2001) Entscheidungsverfahren zur Umsetzung einer nachhaltigen Entwicklung. In: Betriebswirtschaftslehre F (ed) Fakultät Wirtschaftswissenschaften, Technische Universität Dresden

Schulze M (2003) A new monotonic and clone-independent single-winner election method. Voting Matt 17:9–19

Schulze M (2008) A new monotonic, clone-independent, reversal symmetric, and condorcet-consistent single-winner election method (draft)

Seppälä J, Basson L, Norris GA (2002) Decision analysis frameworks for life-cycle impact assessment. J Ind Ecol 5:45–68

Soares RS, Toffoletto L, Deschenes L (2005) Development of weighting factors in the context of LCIA. J Cleaner Prod 14:649–660

Stahl B (1999) Methodenvergleich und Methodenentwicklung zur Lösung der Bewertungsproblematik in produktbezogenen Ökobilanzen. DVS, Digitaler Vervielfältigungs- und Verl.-Service, Frankfurt am Main

Steen BA (2006) Describing values in relation to choices in LCA. Int J Life Cycle Assess 11:277–283

Tangian AS (2000) Unlikelihood of Condorcet's paradox in a large society. Soc Choice Welf 17:337–365

Tideman TN (1985) Investigating the probability of a voting cycle when the electorate is large. Econ Lett 17:23–25

Tideman TN (1987) Independence of clones as a criterion for voting rules. Soc Choice Welf 4:185–206

Vandercruyssen D (1999) Analysis of voting procedures in one-seat elections—Condorcet efficiency and Borda efficiency. In: Studies CfE (ed) Department of Economics, Katholieke Universiteit Leuven

Vetschera R (1991) Entscheidungsunterstützende Systeme für Gruppen—ein rückkopplungsorientierter Ansatz. Physica-Verlag, Heidelberg

Yusoff S, Hansen SB (2007) Feasibility study of performing a life cycle assessment on crude palm oil production in Malaysia. Int J Life Cycle Assess 12:50–58

Zavist TM, Tideman TN (1989) Complete independence of clones in the ranked pairs rule. Soc Choice Welf 6:167–173

Zhang K (2004) Entwicklung eines integrierten multikriteriellen Gruppenentscheidungsunterstützungssystems (MGDSS). Shaker Verlag, Aachen